

MAREK STACHOWSKI
Jagiellonian University in Kraków
marek.stachowski@uj.edu.pl

SLAVIC LANGUAGES IN CONTACT, 1: TOBACCO – SOME SLAVIC DESIGNATIONS IN THEIR AREAL CONTEXT¹

Keywords: Slavic languages, Turkish language, etymology, areal linguistics, lexicology

Abstract

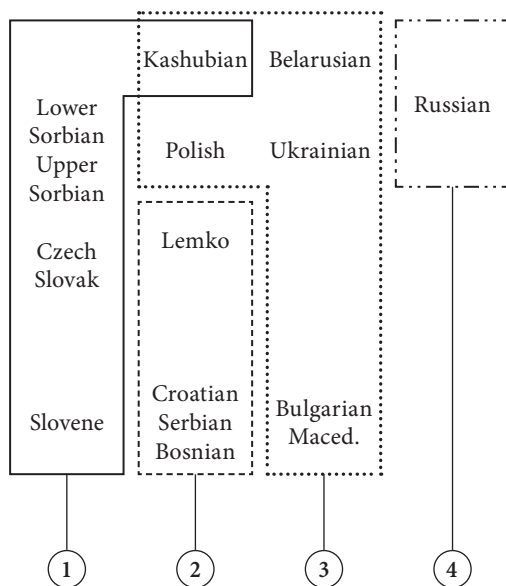
Tobacco was certainly one of the most popular drug products in Europe until the end of the 20th century. Nevertheless, our knowledge of the history of its designations is fairly often limited to general statements like, for instance, “from Spanish” (for English *tobacco*) or “from Turkish” (for Polish *tytoń*) and so on. This author aims at establishing some lexicological areas displaying various influences on the Slavic languages in that respect, as well as providing critical assessment of earlier claims and presenting his own observations concerning the Slavic designations of ‘tobacco’.

Even though I am going to discuss word histories in this study my main point of departure is still the modern areal distribution of literary names of tobacco. Remarks or mentions concerning untypical, dialectal or ephemeral uses and guises are mostly left out here because they certainly deserve separate consideration. To give an example, I can mention the word *tutùm*, attested in 19th century colloquial Dalmatia Croatian and Bosnian instead of *duhan* (Linde 1859: 758). I will also leave aside the fact that, in 17th and 18th century Slovak, the word *dohán* and its phonetic variants

¹ This study was first presented at a conference “Słowianie w monarchii Habsburgów – Literatura, język, kultura” [= ‘Slavs in the Habsburg Monarchy – Literature, Language, Culture’] (June 25th-27th, 2018; Nowy Sącz, The Museum “Miasteczko Galicyjskie”). I was very happy to learn the Lemko form *dùhan* in a discussion following my presentation. – The stress is marked in this paper by grave accent rather than by acute one which makes it possible to clearly distinguish between stressed (e.g. *ù*) and long (e.g. *ú*) vowels in various Slavic languages.

dohán and *dúhan* were used for ‘tobacco’ (HSSJ 1: 275), alongside with *tabak*, because all the *dohán* forms display the then temporary influence of Hungarian upon Slovak.² – See also section “Area 2”.

The modern Slavic designations of tobacco can be divided into groups representing four areas. Area 4 should be considered on its own although one is at first inclined to connect it with area 1. The geographical distribution of the areas is shown in a diagram.



Area 1

The main representatives of this area are the following words (from the north to the south):

- Kashubian *tobak(a)* ‘1. tobacco; 2. snuff’ (according to Internet dictionaries; the word is missing in SEK);
- Lower Sorbian *tubak* (DNW);
- Upper Sorbian *tobak* ~ *dubak* (Schuster-Šewc 1985–1988: 1510);
- Czech *tabák*;
- Slovak *tabak*;
- Slovene *tobak*.

² I do not discuss their plausible but uncertain connection with Czech and Slovak *dochan* ‘pennisetum, fountaingrass’ either. However, an explanation like “afr.” = ‘from the African languages’ in dictionaries of foreign words (SCS 212; ASCS 174) seems rather little informative, especially in light of the number and diversity of the African languages. It is particularly the adjective *akademický* in the title of ASCA that seems to be an indication of a somewhat more precise etymology here.

The Kashubian form *tobaka* most probably results from the contamination of Kashubian (< German) *Toback* ‘tobacco’ with Polish *tabaka* ‘snuff’.

The *u* variants (in Lower Sorbian and partially Upper Sorbian) are clearly secondary and younger than the *o* ones; they result from inner Sorbian evolution.

The origin of the *o* variants can easily be explained: they are all Slavic reflexes of the older German shape *Toback* (< French *tobac*) id. (Schuster-Šewc 1985–1988: 1510). Cf. also “Ältere und noch mundartl. Nebenform *Tobak* (engl. *tobacco*); vgl. die entsprechende Doppelformigkeit bei *Kaffee* – *Koffee*. Umgangsspr. *Anno Tobak* ‘vor langer Zeit’ (umgestaltet aus *Anno Domini*)” (Paul 1961: 611). On the other hand, however, the Slovene variant *tobak* seems to be recorded in the 18th century whereas *tabak* was first attested in the 17th century (Bezljaj 2005: 188). That fact warrants explanation.³ In purely phonetic terms also the older French variant *tobac* could be considered an etymon of the Slavic words but the history of the language contacts in East Europe clearly gives preference to the German source.

The Czech and the Slovak words have an *-a-* rather than *-o-* in the first syllable. So one may say that both words reflect the newer German form *Tabak*. It is, however, not clear why Czech and Slovak should have borrowed the German designation of tobacco later than Slovene, Sorbian and Kashubian did. Also the geographical position of Czech and Slovak between the Slovene *to-* variant, on the one hand, and the Sorbian *to-* form, on the other, makes a possibility of the influence of quite a different German dialect less convincing. The origin of the *o* ~ *a* difference in those Slavic words does not seem to have ever been discussed and still less established in etymological dictionaries.

The long vowel of Czech *tabák* does not fit the German initial stress of *Tabak*. Both that fact and the lack of a long vowel in its Slovak counterpart remain unexplained, possibly even undiscussed. It is therefore instructive to mention the following explanation of the situation in German dialects: “*Tabak*. Jetzt nordd. *Tábak* betont. Die ältere noch südd. Betonung ist in den Zuss. *Kaut.*, *Raucht.*, *Schnupft.* allgemein geblieben, in denen sie der Sprachgewohnheit bequemer liegt” (Paul 1961: 611).⁴ That points to a possible South German source of Czech *tabák*, whereas Slovak *tabak* should probably rather be viewed as a reflex of the general literary German form *Tabak* with its initial stress. Another possibility was suggested by Vít Boček (Brno; personal communication) who believes that a German loanword **tabak* was adapted to genuine Czech *-ák* derivatives in the process of folk etymology (like *kuřák* ‘smoker’, *tesák* ‘bowie-knife’, *levák* ‘left-handed person’, and so on). I would also like to add that this interpretation very well explains both the absence of a long *-á-* in Slovak (folk etymology and levels of morphological adaptation are psychologically guided phenomena and one cannot expect them to work in parallel and consistently even

³ Interpretation of chronological discrepancies often involves additional queries and problems. After all, the difference between December 31st, 1700 and January 1st, 1701, too, is that between the 17th and the 18th century.

⁴ Cf. also the German phonetic dictionary: “**Tabak** ‘ta:bak, auch ta’bak” (WDA 483). The case is not unique among the loanwords in German, cf. similarly “**Orient** ‘o:rient od. o’rient” (WDA 389); “**Suffix** ‘zʊfiks od. zʊ’fiks” (WDA 479) as well as “*Wallach* (älter *Wallách*)” (Bielfeldt 1965: 19).

though the suffix *-ák* shows a long vowel in Slovak, too) and the fact that Czech derivatives of *tabák* (like the adjective *tabákový*) display a long vowel, whereas Czech loanwords reflecting derivatives created in other languages do not, good examples being medical terms like *tabakizmus* and *tabakóza* (ASCS 741).

In this respect, deserving of attention is a verbal collocation with the literal meaning ‘to drink tobacco’ (that is ‘to smoke [tobacco, cigarette, pipe]’). The Slovene expression *piti tobak* is being traced back to a northern source like German *Tabak trinken*, Old English *drink tobacco*, Danish *drikke tobak*, and the like, in Bezljaj’s etymological dictionary (Bezljaj 2005: 188). I do not see why exactly those languages were accepted by France Bezljaj (only the German influence upon Slovene in the past centuries warrants mention in that context) even though he was perfectly aware of their South European counterparts like Serbian dialectal *pijem tutun*, Albanian *pi duhan* or Rumanian *a bea tutun* (Bezljaj 2005: 188). All the Balkan examples are usually viewed as reflexes of the Turkish expression *tütün içmek* id. which is connected with the use of water pipes. Here, tobacco is “drunk” because the smoke is filtered through water. I would rather therefore accept the Turkish collocation as a source of both Balkanic and Germanic data. The exact channel of transmission should only be examined in the future, and collaboration of linguists with cultural historians will decidedly be necessary. Bezljaj (2005: 188) thinks rather of a calque of Latin (*hauritur, bibitur*) and Greek (*καταπίνεται*) medical terms which is open to doubt. Additionally, the Italian idiom *fumare come un turco* makes me think that the Turks were perceived in Europe as a symbol of smoking and, thus, the Germanic expressions will presumably have come into being as calques of the Turkish collocation rather than otherwise. The way things are at the moment, a detailed explanation of the mutual relation between the Latin, Greek, Germanic, Balkan and Turkish collocations still looms on the horizon, the more so as the use of a water filter was also known in non-Balkan Europe (cf. Černych 1993: 223b).

Area 2

The Croatian, Serbian and Bosnian languages display the same word for ‘tobacco’, namely *duhan*, as well as, more or less sporadically, its phonetic variant: *duvan*. Though ranked 565, that colloquial form is marked as a non-standard one in the Croatian frequency list, while the standard form *duhan* is ranked 508 (HČR 335) – the difference does not appear significant.

The Lemko variety of Ruthenian⁵ is another language in which the word *duhan* is being used. This fact seems to be especially important in our context, as will be shown below.

⁵ Semantic limitation of the term *Ruthenian* to East Slavs living “in the Grand Duchy of Lithuania and the Polish Crown in contradistinction to *Moscovitae*” was introduced by the 16th century (Danylenko 2017: 166, fn. 1) and this term seems to be preferred to the name (*Lemko-*)*Rusyn* by some Lemkos even today.

The direct etymon of the South Slavic variant is generally considered to be Turkish *duhan* ‘1. smoke, fume; 2. tobacco’ that, in its turn, reflects Arabic *duḥān* ‘smoke, fume’ (Lokotsch 1927: 43; S. Stachowski 1975: 49). Another trace of that Arabo-Turkish word is Hungarian *dohány* ‘tobacco’.

The problem is why all these languages borrowed *duhan* rather than *tütün* id. (see “Area 3” below) even though the latter word has been much more generally used for ‘tobacco’ in Turkish.

The Lemko attestation seems to suggest a reasonable solution: The Turkish word *duhan* did not spread from Turkish over Bosnian etc. northwards to Lemko but it was, instead, originally borrowed into Hungarian and only then transmitted from Hungarian into both Lemko in the north and the three Slavic languages in the south. Sure, one can hardly explain why *duhan* rather than *tütün* was borrowed by the Hungarians (should this word choice have been connected with a sort of smoking material?). Nevertheless, that choice seems more plausible for Hungarian alone rather than for five languages in the vast territory between Herzegovina and Poland.

Thus, the word *duhan* in Lemko,⁶ Croatian, Serbian as well as in Bosnian should be viewed as a reflex of Hungarian *dohány* (< Turkish *duhan* < Arabic *duḥān*).

For the situation in the past centuries cf. also the older Slovak variant *doháň* mentioned in the introductory part of this article.

Area 3

This group contains European reflexes of the Turkish word *tütün* ‘1. smoke, fume; 2. tobacco’ (< *tüt-* ‘to emit smoke’):

- Polish *tytoń* (= older Polish *tytuń* ~ *tutuń* ~ *tiutuń* ~ *tiutiun* id. [Linde 1859: 758]);
- Belarusian *tytun* (~ *tytun* ~ *cjucjun* [ESUM 5: 696]);
- Ukrainian *tjutjun* (~ *tjutin* [ESUM 5: 695]);
- Bulgarian *tjutjun*;
- Macedonian *tutun* (in relatively recent years also *duan* [Jašar-Nasteva 2001: 101] – possibly a loanword from Bosnian *duhan*; not attested in Bitola [Cvetkovski 2017])

As is generally the case with loanwords, an ultimate etymon can be ascertained more easily than the borrowing channels. The palatality of the Slavic word-final *-ń* (< *-n*) could have been explained as the result of the Hungarian impact because that language has only the palatal *-ń* which makes the adaptation of *-n* in the borrowings difficult. However, a word matching the Polish and the Belarusian data is not found in Hungarian. Thus, the Slavic *-ń* should probably be understood as the effect of palatal assimilation and dissimilation of the middle and the word-final consonants, that is: **t' – n > t' – ń > t – ń*, for instance in Polish: **tiutiun > tiutiun > tiutuń* (and, then, further: *> tytuń > tytoń*).

⁶ NB, note the word-initial stress in Lemko *dūhan*, exactly rendering the Hungarian prosodical pattern, as opposed to the ultima stress in Turkish.

The Belarusian forms with *ty-* presumably go back to older Polish *tytuń*. If this conjecture is correct, the stress of the Belarusian *tytùń* suggests a contamination of Polish *tytuń* with its paroxytonic stress and Ukrainian or Russian (see Area 4) *tjutjùn* id.

As far as the older and dialectal Polish form *tytuń* is concerned, the greatest problem is its vowel sequence *-y – u-*. Actually there is nothing strange about non-Slavic *ü* being rendered by *y* when borrowed into a Slavic language (e.g. Polish *dybel* ‘wall plug’ < German *Dübel* id.; NB, compare also the Polish-German rules of using the characters ‘y’ and ‘ü’ in Franciscus Meninski’s Ottoman-Turkish dictionary [M. Stachowski 2012: 195, 196]). However, it is not really clear why Turkish *tütün* was rendered by *tytuń* rather than by **tytyń* or just **tytyn* in Polish. If the chain of palatal changes, as shown above, is correct the only really attested form preceding *tytuń* is *tiutuń* which points to a somewhat peculiar change of *-’u-* > *-y-*. Arguably, that would be only possible if a speaker identified the Slavic *-’u-* with the Turkish *-ü-* which, however, is difficult to imagine. One can assume that the variant *tutuń* seemed to have had more chances to take roots in Polish.

Finally, we also have to reckon with a possibility of two different sources, namely, a Kipchak one for the northern languages in this area, and a Turkish one for Bulgarian and Macedonian.

The purely linguistic methods seem to fail in more detailed investigations. Philological examination of mercantile texts, linked with historical insights will hopefully help us show the routes taken by the Turkish word *tütün* in the Slavic world which, however, is a task for the future.

Area 4

This is the smallest group in our study – one consisting of Russian *tabák* ‘tobacco’ only (the older and dialectal form *tjutjùn* must be excluded in our context). Černych (1993: 223) is right about attaching much significance to the question of stress. The Russian word could hardly be borrowed from Polish *tabaka* ‘snuff’. But Černych’s claim that also German *Tabak* has to be excluded because of the German initial stress is only partially justified – as mentioned above, the older South German form *Tabak* originally had an oxytone accent and, thus, it could theoretically have been a source of the Russian word. In terms of cultural history, however, French *tabac* seems an equally good or, maybe, even better candidate though it is pronounced without the final <-c> because tobacco might have first become known to the Russian intelligentsia through texts written in French. Černych (1993: 223) cursorily mentions the Dutch variant *tabak*, stressed on the ultima. This suggestion, missing in Fasmer (1987: 5), appears very interesting because of Peter the Great’s years in the Netherlands and his predilection for the Dutch culture and language (cf. the Dutch origin of the name of the Tsarist capital Saint Petersburg, originally, even though only temporarily, called à la hollandaise *Санкт-Питербурх* (for *Sankt-Pieterboerg*); note also that the modern colloquial short form still is *Питер* rather than **Пемеп*).

If this conjecture is correct, Czech *tabák* and Slovak *tabak* are German loanwords, whereas Russian *tabák* is a Dutch one.

*
* *

It is easy to note that the originally American word *tobacco* actually only arrived at the western fringe of the Slavic world (the Russian *tabák* should rather be excluded here if it actually is an artificial innovation resulting from the intentional promotion policy of Dutch).

Since Turkish *tütün* and *duhan* have at first just meant 'smoke, fume' they are not directly connected with either cultivated (*Nicotiana tabacum*, *Nicotiana rustica*) or one of "some five dozen wild species of *Nicotiana*, three-quarters of them native to the Americas" (Pendell 2010: 31). They were initially, that is before Turks, Arabs and other peoples living in the Middle East became familiar with the American tobacco, presumably used to denote a herbal and/or fruit blend or, maybe, cannabis used in water pipes.⁷ A question of both cultural and linguistic character is whether the words *tütün* and *duhan* were used for non-tobacco or cannabis blends in the 15th and 16th century Balkans and always with the same meaning.

*
* *

In sum, a few observations were made and a few phenomena could be commented upon here. The five facts observed and better (or for the first time at all) discussed but still not solved are the following: (1) Czech and Slovak *ta-*, not *to-*; (2) Slovene chronology of *tobak* and *tabak*; (3) the origin and channels of transmission of the collocation 'to drink tobacco'; (4) the possibility of Hungarian influence on Croatian, Serbian and Bosniak *duhan* as well as Lemko *dùhan*; (5) Polish *tytuń* in lieu of **tytyń* or **tytyn*.

Problem (6) remains a focus of the experts in the local history of the Balkan lands: The meaning of *tütün* and *duhan* before the American tobacco became popular in the Levant (and then in the Balkans) remains a question to be answered by cultural historians rather than by linguists.

The following five problems can be viewed as resolved or at least discussed and better rendered than it was the case earlier: (A) the origin of Kashubian *tobaka*; (B) the origin of the Czech long *-á-* in contrast to the Slovak short *-a-*; (C) long *-á-* in Czech *tabákový* vs. short *-a-* in Czech *tabakizmus ~ tabakóza*; (D) the origin of *-ń* in Polish; (E) the origin of Belarusian *tytùń*; (E) the origin of Russian *tabák*.

*
* *

If my argumentation and the borrowing routes as suggested in the foregoing analysis are correct the four areas represent four direct borrowing sources: Area 1 – German; Area 2 – Hungarian; Area 3 – Turkish; Area 4 – Dutch.

⁷ For the use of cannabis, also in usual pipes, in Asia and Russia see K. Stachowski 2004, continued in Alonso de la Fuente 2007.

References

- ASCS = Petráčková V., Krause J. et al. 1998. *Akademický slovník cizích slov*. Praha.
- DNW = Starosta M., Hannusch E., Bartels H. n.d. *Deutsch-niedersorbisches Wörterbuch*. [www.dolnoserbski.de].
- ESUM = Mel'nychuk O.S. et al. (eds.) 2006. *Etymologičnyj slovnyk ukrajins'koji movy*. [vol. 5]. Kyjiv.
- HČR = Moguš M., Bratanić M., Tadić M. 1999. *Hrvatski čestotni rječnik*. Zagreb.
- HSSJ = Majtán M. (ed.) 1991. *Historický slovník slovenského jazyka*. [vol. 1]. Bratislava.
- SCS = Petráčková V., Krause J. et al. 1997. *Slovník cudzích slov*. Bratislava.
- SEK = Boryś W., Popowska-Taborska H. 1994–2010. *Słownik etymologiczny kaszubszczyzny*. [vol. 1–6]. Warszawa.
- WDA = Stötzer U. et al. 1974. *Wörterbuch der deutschen Aussprache*. Leipzig.
- Alonso de la Fuente J.A. 2007. How far do pipes travel? – *Central Asiatic Journal* 51.2: 161–166.
- Bielfeldt H.H. 1965. Die Entlehnungen aus den verschiedenen slavischen Sprachen im Wortschatz der neuhochdeutschen Schriftsprache. – *Sitzungsberichte der Deutschen Akademie der Wissenschaften Berlin, Klasse für Sprache, Literatur und Kunst* 1965.1: 3–59.
- Bezljaj F. 2005. *Etimološki slovar slovenskega jezika*. [vol. 4]. Ljubljana.
- Černych P.Ja. 1993. *Istoriko-ëtimologičeskij slovar' sovremennogo russkogo jazyka*. [vol. 2]. Moskva.
- Cvetkovski V. 2017. *Etnolingvistički osobenosti na turskite leksički zaemki vo kolokvijalniot makedonski jazik vo Bitola*. Skopje.
- Danylenko A. 2017. On the names of Ruthenia in early modern Poland-Lithuania. – Németh M., Podolak B., Urban M. (eds.): *Essays in the history of languages and linguistics*. Kraków: 161–173.
- Fasmer [= Vasmer] M. 1987. *Ëtimologičeskij slovar' russkogo jazyka*. [vol. 4]. Moskva.
- Jašar-Nasteva O. 2001. *Turskite leksički elementi vo makedonskiot jazik*. Skopje.
- Linde S.B. 1859. *Słownik języka polskiego*. [vol. 5]. Lwów.
- Lokotsch K. 1927. *Etymologisches Wörterbuch der europäischen (germanischen, romanischen und slavischen) Wörter orientalischen Ursprungs*. Heidelberg.
- Paul H. 1961 [ed. by A. Schirmer]. *Deutsches Wörterbuch*. Halle (Saale).
- Pendell D. 2010. *Pharmako/Poeia: Plant powers, poisons, and herbcraft*. Berkeley.
- Schuster-Šewc H. 1985–1988. *Historisch-etymologisches Wörterbuch der ober- und niedersorbischen Sprache*. [vol. 3]. Bautzen.
- Stachowski K. 2004. A Siberian word for 'pipe' and its possible Indo-European cognates. – *Studia Etymologica Cracoviensia* 9: 125–132.
- Stachowski M. 2012. Remarks on the phonetic value of the letters ⟨y⟩ and ⟨ü⟩ in Franciscus Meninski's Ottoman Turkish *Thesaurus* (1680). – *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 129.3: 189–197.
- Stachowski S. 1975. *Studien über die arabischen Lehnwörter im Osmanisch-Türkischen*. [vol. 1]. Wrocław, Warszawa, Kraków, Gdańsk.